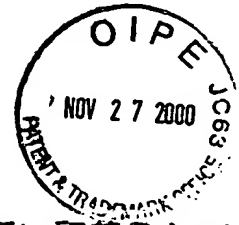


500P0982US00

日 本 国 特 許 庁  
PATENT OFFICE  
JAPANESE GOVERNMENT



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

1999年 8月24日

出 願 番 号

Application Number:

平成11年特許願第236412号

出 願 人

Applicant (s):

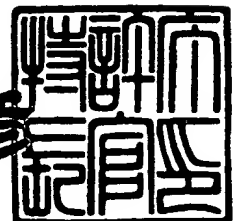
ソニー株式会社

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2000年 6月23日

特許庁長官  
Commissioner,  
Patent Office

近 藤 隆 彦



出証番号 出証特2000-3047870

【書類名】 特許願

【整理番号】 9900232202

【提出日】 平成11年 8月24日

【あて先】 特許庁長官殿

【国際特許分類】 G11B 20/00

【発明者】

【住所又は居所】 東京都品川区北品川6丁目7番35号 ソニー株式会社  
内

【氏名】 水藤 太郎

【発明者】

【住所又は居所】 東京都品川区北品川6丁目7番35号 ソニー株式会社  
内

【氏名】 太田 正志

【特許出願人】

【識別番号】 000002185

【氏名又は名称】 ソニー株式会社

【代表者】 出井 伸之

【代理人】

【識別番号】 100091546

【弁理士】

【氏名又は名称】 佐藤 正美

【電話番号】 03-5386-1775

【手数料の表示】

【予納台帳番号】 048851

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

●  
特平 1 1 - 2 3 6 4 1 2

【包括委任状番号】 9710846

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声信号再生方法および音声信号再生装置

【特許請求の範囲】

【請求項 1】

記録媒体から、これに記録された音声信号を定常速度より速い速度で再生することによって得られた再生入力音声信号を、処理単位期間ごとに区切り、

出力バッファから定常速度相当量の再生出力音声信号が得られる範囲内で無音部分を削除し、前後の有音部分を接合して、それぞれの処理単位期間の再生入力音声信号を、前記出力バッファに書き込むとともに、

処理単位期間の再生入力音声信号中に前記出力バッファに蓄えられない量の有音部分が存在する場合には、一部の信号部分を削除し、前後の有音部分を接合して、その処理単位期間の再生入力音声信号を圧縮して、前記出力バッファに書き込む音声信号再生方法。

【請求項 2】

請求項 1 の音声信号再生方法において、

無音部分を削除し、前後の有音部分を接合するに当たっては、接合点直前の有音部分をフェードアウト処理し、接合点直後の有音部分をフェードイン処理することを特徴とする音声信号再生方法。

【請求項 3】

請求項 1 または 2 の音声信号再生方法において、

処理単位期間の再生入力音声信号を圧縮するに当たっては、フェード期間において前後の有音部分を重ね合わせるとともに、そのフェード期間における前後の有音部分の差分が最小となるように前後の有音部分の位置関係を調整した上で、クロスフェードによって前後の有音部分を接合することを特徴とする音声信号再生方法。

【請求項 4】

請求項 1 ～ 3 のいずれかの音声信号再生方法において、

処理単位期間を固定の時間とすることを特徴とする音声信号再生方法。

【請求項 5】

請求項 1 ～ 3 のいずれかの音声信号再生方法において、  
再生入力音声信号の主成分の音程に応じて処理単位期間の時間を変化させることを特徴とする音声信号再生方法。

【請求項 6】

請求項 1 ～ 3 のいずれかの音声信号再生方法において、  
再生入力音声信号の無音部分と有音部分の境目、または有音部分中の相対的にレベルが低い部分を、処理単位期間の区切れ目とすることを特徴とする音声信号再生方法。

【請求項 7】

記録媒体から、これに記録された音声信号を定常速度より速い速度で再生することによって得られた再生入力音声信号に対する速度変換処理の処理単位期間を設定する処理単位期間設定部と、

再生出力音声信号を得るための出力バッファと、

この出力バッファから定常速度相当量の再生出力音声信号が得られる範囲内で無音部分を削除し、前後の有音部分を接合して、それぞれの処理単位期間の再生入力音声信号を、前記出力バッファに書き込むとともに、処理単位期間の再生入力音声信号中に前記出力バッファに蓄えられない量の有音部分が存在する場合には、一部の信号部分を削除し、前後の有音部分を接合して、その処理単位期間の再生入力音声信号を圧縮して、前記出力バッファに書き込む速度変換処理部と、  
を備える音声信号再生装置。

【請求項 8】

請求項 7 の音声信号再生装置において、

前記速度変換処理部は、無音部分を削除し、前後の有音部分を接合するに当たって、接合点直前の有音部分をフェードアウト処理し、接合点直後の有音部分をフェードイン処理することを特徴とする音声信号再生装置。

【請求項 9】

請求項 7 または 8 の音声信号再生装置において、

前記速度変換処理部は、処理単位期間の再生入力音声信号を圧縮するに当たっ

て、フェード期間において前後の有音部分を重ね合わせるとともに、そのフェード期間における前後の有音部分の差分が最小となるように前後の有音部分の位置関係を調整した上で、クロスフェードによって前後の有音部分を接合することを特徴とする音声信号再生装置。

【請求項 1 0】

請求項 7～9 のいずれかの音声信号再生装置において、

前記処理単位期間設定部は、処理単位期間を固定の時間とすることを特徴とする音声信号再生装置。

【請求項 1 1】

請求項 7～9 のいずれかの音声信号再生装置において、

当該音声信号再生装置は、さらに再生入力音声信号の主成分の音程を検出する音程検出部を備え、

前記処理単位期間設定部は、この音程検出部の検出結果に基づいて、再生入力音声信号の主成分の音程に応じて処理単位期間の時間を変化させることを特徴とする音声信号再生装置。

【請求項 1 2】

請求項 7～9 のいずれかの音声信号再生装置において、

前記処理単位期間設定部は、再生入力音声信号の無音部分と有音部分の境目、または有音部分中の相対的にレベルが低い部分を、処理単位期間の区切れ目とすることを特徴とする音声信号再生装置。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

この発明は、ビデオテープレコーダなどのように映像信号および音声信号を記録再生する機器や、ミニディスクプレーヤなどのように音声信号を記録再生する機器などにおいて、記録媒体から、これに記録された音声信号を定常速度より速い速度で再生する方法および装置に関する。

【0 0 0 2】

【従来の技術】

家庭用ビデオテープレコーダで、テープに記録された映像信号および音声信号を、定常速度より速い2倍速などの速度で再生する場合でも、音声信号については、無音部分を優先的に削除するように元の再生音声信号を周期的に切り貼りすることによって、音声を定常速度で聞き取ることができるようにしたものが考えられている。

【0 0 0 3】

【発明が解決しようとする課題】

しかしながら、上述した従来の音声信号再生方法では、元の再生音声信号を周期的に削除するため、元の再生音声信号中に無音部分が一定の割合で含まれていない場合には、有音部分が削除されて、切り貼り後の再生音声信号は音声途切れ途切れになってしまい、内容を理解することができなくなってしまう。

【0 0 0 4】

そこで、この発明は、第1に、記録媒体から、これに記録された音声信号を定常速度より速い速度で再生し、定常速度に変換して出力する方法において、音声途切れず、内容を理解でき、聴きやすいようにしたものである。

【0 0 0 5】

この発明は、第2に、上記の第1に加えて、ノイズなどの少ない、音質の良い出力音声信号が得られるようにしたものである。

【0 0 0 6】

【課題を解決するための手段】

この発明の音声信号再生方法では、

記録媒体から、これに記録された音声信号を定常速度より速い速度で再生することによって得られた再生入力音声信号を、処理単位期間ごとに区切り、

出力バッファから定常速度相当量の再生出力音声信号が得られる範囲内で無音部分を削除し、前後の有音部分を接合して、それぞれの処理単位期間の再生入力音声信号を、前記出力バッファに書き込むとともに、

処理単位期間の再生入力音声信号中に前記出力バッファに蓄えられない量の有

音部分が存在する場合には、一部の信号部分を削除し、前後の有音部分を接合して、その処理単位期間の再生入力音声信号を圧縮して、前記出力バッファに書き込む。

## 【 0 0 0 7 】

上記の方法による、この発明の音声信号再生方法では、出力バッファを効果的に用いることによって、必要最小限のメモリを用いるだけで、音声が続切れず、内容を理解でき、聴きやすい状態で、定常速度より速い速度で再生された音声信号を定常速度に変換して出力することができる。

## 【 0 0 0 8 】

## 【発明の実施の形態】

## 〔システム全体の一実施形態…図 1〕

図 1 は、この発明の再生装置の一実施形態を示し、ハードディスク、光磁気ディスク、光ディスクなど、ディスク状の記録媒体から、これに記録された映像信号および音声信号を、定常速度より速い速度で再生することができる場合である。

## 【 0 0 0 9 】

再生装置は、記録装置を兼ねてもよく、実際上も記録装置を兼ねることが望ましいが、記録方法および記録装置は、公知のものと特に変わらないので省略し、記録媒体 1 には、所定のプロトコルによるデジタル化された映像信号および音声信号が記録されているものとする。音声信号のサンプリング周波数は、例えば 48 k H z である。

## 【 0 0 1 0 】

記録媒体 1 は、駆動モータ 3 によって回転駆動される。再生ヘッド 2 は、記録媒体 1 から、これに記録されている映像信号および音声信号を読み取るもので、磁気ヘッドや光学ピックアップなどであり、駆動モータ 4 を含む移動機構によって記録媒体 1 の径方向に移動させられる。駆動モータ 3 および 4 は、サーボ回路 5 によって駆動制御され、サーボ回路 5 は、装置全体を制御するシステムコントローラ 6 によって制御される。



【0 0 1 1】

再生ヘッド2の出力、すなわち記録媒体1から読み取られた映像信号および音声信号は、倍速デコーダ7に供給され、倍速デコーダ7において、映像信号が、再生倍率に応じてフレーム単位またはフィールド単位で間引かれるとともに、音声信号が、再生倍率にかかわらず上記のサンプリング周波数が維持されるようにサンプル単位で間引かれる。

【0 0 1 2】

すなわち、2倍速再生の場合には、記録媒体1からは映像信号および音声信号が定常速度の2倍の速度で読み取られるが、倍速デコーダ7において、記録媒体1から読み取られた映像信号が、連続する2フレームにつき1フレーム、または連続する2フィールドにつき1フィールド、間引かれるとともに、記録媒体1から読み取られた音声信号が、連続する2サンプルにつき1サンプル、間引かれる。

【0 0 1 3】

3倍速再生の場合には、記録媒体1からは映像信号および音声信号が定常速度の3倍の速度で読み取られるが、倍速デコーダ7において、記録媒体1から読み取られた映像信号が、連続する3フレームにつき2フレーム、または連続する3フィールドにつき2フィールド、間引かれるとともに、記録媒体1から読み取られた音声信号が、連続する3サンプルにつき2サンプル、間引かれる。

【0 0 1 4】

したがって、音声信号のサンプリング周波数が48kHzで、60m秒の時間では2880サンプルの場合、再生倍率にかかわらず、倍速デコーダ7の出力の音声信号は、60m秒の時間では2880サンプルとなる。

【0 0 1 5】

倍速デコーダ7の出力の映像信号は、映像処理部8で記録時の圧縮に対応して伸長され、アナログ信号に変換されて、CRTディスプレイや液晶ディスプレイなどの表示装置10に供給される。

【0 0 1 6】

倍速デコーダ7の出力の音声信号は、音声処理部9で後述のように処理され、

アナログ信号に変換されて、スピーカなどの音声出力装置 1 1 に供給される。

【0 0 1 7】

システムコントローラ 6 には、操作部 1 2 が接続される。操作部 1 2 は、装置の利用者が再生倍率を指示するなどの各種操作を行うもので、これには、装置の動作状態や操作状況を表示する液晶表示部などの表示部が設けられる。

【0 0 1 8】

以下に、音声処理部 9 の構成および動作を示す。以下の処理は 2 倍速再生の場合を例として示すが、3 倍速再生などの場合にも同様の処理が行われる。

【0 0 1 9】

〔音声処理部の一例と速度変換処理…図 2 ～図 7〕

図 2 は、音声処理部 9 の一例を示す。この例は、音声信号についての速度変換処理の処理単位時間を一定にする場合で、音声処理部 9 は、処理単位期間設定部 2 1、有音無音判定部 2 2、速度変換処理部 2 3、出力バッファ 2 4、D/A コンバータ 2 5 および音声増幅器 2 6 によって構成される。

【0 0 2 0】

処理単位期間設定部 2 1 は、図 1 のシステムコントローラ 6 からのタイミング信号およびクロックによって、図 3 に処理単位期間 T 1 ～T 6 として示すように、一音素の時間より短い 6 0 m 秒の時間間隔で、速度変換処理のそれぞれの処理単位期間の先頭ないし末尾のタイミングを決定する。ただし、処理単位期間設定部 2 1 は、システムコントローラ 6 の一部、または速度変換処理部 2 3 の一部として、構成することもできる。

【0 0 2 1】

倍速デコーダ 7 の出力の音声信号、すなわち図 3 に入力データ S i として示す音声データは、それぞれの処理単位期間では、上述したように 2 8 8 0 サンプルである。なお、図 3 以下に入力データ S i などとして示す音声データは、いずれもデジタル信号であるが、発明の理解を容易にするため、アナログ信号に変換された後の波形で示す。

【0 0 2 2】

有音無音判定部 2 2 は、入力データ S i、すなわち倍速デコーダ 7 からの音声

信号の、有音部分と無音部分を識別判定するもので、この例では、図 4 に示すように、入力データ  $S_i$  の、それぞれ 60 m 秒、2880 サンプルの処理単位期間を、3 m 秒、144 サンプルずつの 20 個の判定単位期間に分割し、それぞれの判定単位期間につき、144 サンプルのデータの絶対値の平均値を算出して、その平均値が閾値を超えるか否かを判断し、平均値が閾値を超えるときには、その判定単位期間を有音部分と判定し、平均値が閾値以下であるときには、その判定単位期間を無音部分と判定する。

## 【0023】

出力バッファ 24 は、FIFO (First In First Out) 構造のリングバッファとして機能するものとする。図 5 に、出力バッファ 24 へのデータ書き込み、出力バッファ 24 からのデータ読み出し、および出力バッファ 24 の入出力ポインタの動きを示す。

## 【0024】

出力バッファ 24 にデータが全く書き込まれていないときには、図 5 (A) に示すように、出力バッファ 24 には先頭アドレスから順にデータが書き込まれ、それに応じて入力ポインタも後方に移動する。このとき、出力ポインタは先頭アドレスを指示する。

## 【0025】

図 5 (B) は、この状態から、さらにデータが書き込まれるとともに、先頭アドレスから順にデータが読み出された状態を示し、出力ポインタも後方に移動する。

## 【0026】

さらにデータが書き込まれて、入力ポインタが末尾アドレスに達すると、図 5 (C) に示すように、入力ポインタは先頭アドレスに戻る。このとき、新たなデータが出力済みデータに対してのみ上書きされるように、入力ポインタの指示位置が制御される。出力ポインタも、末尾アドレスに達したときには、先頭アドレスに戻る。

## 【0027】

D/A コンバータ 25 は、出力バッファ 24 から出力された音声データをアナ

ログ音声信号に変換するものであり、音声増幅器 2 6 は、そのアナログ音声信号を増幅するものである。音声増幅器 2 6 からのアナログ音声信号は、図 1 の音声出力装置 1 1 に供給される。

【 0 0 2 8 】

速度変換処理部 2 3 は、図 6 に示す速度変換処理ルーチンによって、倍速デコーダ 7 の出力の音声信号、すなわち入力データ  $S_i$  を速度変換するものである。

【 0 0 2 9 】

すなわち、この速度変換処理ルーチン 3 0 では、まずステップ 3 1 において、有音無音判定部 2 2 の判定結果から、入力データ  $S_i$  のそれぞれの処理単位期間内に無音部分が倍率相当分以上存在するか否かを判断する。倍率相当分とは、処理単位時間に再生倍率の逆数を乗じた時間で、処理単位時間が 6 0 m 秒で、かつ 2 倍速再生の場合には、3 0 m 秒分である。すなわち、ステップ 3 1 では、それぞれの処理単位期間内に無音部分が 3 0 m 秒分（1 4 4 0 サンプル）以上存在するか否かを判断する。

【 0 0 3 0 】

そして、処理単位期間内に無音部分が 3 0 m 秒分（1 4 4 0 サンプル）以上存在するときには、ステップ 3 1 からステップ 3 2 に進んで、出力バッファ 2 4 に 3 0 m 秒分（1 4 4 0 サンプル）以上のデータが蓄積されているか否かを判断し、出力バッファ 2 4 に 3 0 m 秒分（1 4 4 0 サンプル）以上のデータが蓄積されているときには、ステップ 3 2 からステップ 3 3 に進んで、その処理単位期間内の無音部分を全て削除し、無音部分の前後の有音部分を接合して、その処理単位期間内の有音部分のデータのみを全て、出力バッファ 2 4 に出力して書き込む。

【 0 0 3 1 】

また、出力バッファ 2 4 に 3 0 m 秒分（1 4 4 0 サンプル）以上のデータが蓄積されていないときには、ステップ 3 2 からステップ 3 4 に進んで、その処理単位期間内の 3 0 m 秒分（1 4 4 0 サンプル）以下の有音部分の全てに、その処理単位期間の先頭に近い部分から、その処理単位期間内の無音部分を一部付加して、その処理単位期間内の 3 0 m 秒分（1 4 4 0 サンプル）のデータを、出力バッファ 2 4 に出力して書き込む。

【0 0 3 2】

ただし、処理単位期間内に有音部分と無音部分が、ちょうど 3 0 m 秒分（1 4 4 0 サンプル）ずつ存在するときには、ステップ 3 3 と同様に、その処理単位期間内の無音部分を全て削除し、無音部分の前後の有音部分を接合して、その処理単位期間内の有音部分のデータのみを全て、出力バッファ 2 4 に出力して書き込む。

【0 0 3 3】

一方、処理単位期間内に無音部分が 3 0 m 秒分（1 4 4 0 サンプル）未満しか存在しないときには、すなわち有音部分が 3 0 m 秒分（1 4 4 0 サンプル）を超えて存在するときには、ステップ 3 1 からステップ 3 5 に進んで、出力バッファ 2 4 にその処理単位期間内の 3 0 m 秒分（1 4 4 0 サンプル）を超える有音部分のデータを全て蓄えるだけの空きがあるか否かを判断する。

【0 0 3 4】

そして、出力バッファ 2 4 にその処理単位期間内の 3 0 m 秒分（1 4 4 0 サンプル）を超える有音部分のデータを全て蓄えるだけの空きがあるときには、ステップ 3 5 からステップ 3 3 に進んで、その処理単位期間内の無音部分を全て削除し、無音部分の前後の有音部分を接合して、その処理単位期間内の有音部分のデータのみを全て、出力バッファ 2 4 に出力して書き込む。

【0 0 3 5】

ただし、その処理単位期間内の 6 0 m 秒分（2 8 8 0 サンプル）のデータが全て有音部分であるときには、その 6 0 m 秒分（2 8 8 0 サンプル）の有音部分のデータをそのまま、出力バッファ 2 4 に出力して書き込む。

【0 0 3 6】

また、出力バッファ 2 4 にその処理単位期間内の 3 0 m 秒分（1 4 4 0 サンプル）を超える有音部分のデータを全て蓄えるだけの空きがないときには、ステップ 3 5 からステップ 3 6 に進んで、有音部分が 3 0 m 秒分（1 4 4 0 サンプル）となるようにデータを圧縮し、その圧縮後のデータを、出力バッファ 2 4 に出力して書き込む。後述のように、この場合は速聴きとなる。

【 0 0 3 7 】

この圧縮処理について示すと、例えば、図 7 の上段に示すように、その処理単位期間内の 6 0 m 秒分（2 8 8 0 サンプル）のデータが全て有音部分であるときには、同図の下段に示すように、中間部分の 1 0 4 0 サンプルのデータを削除し、期間の先頭から 9 2 0 サンプルの前方部データ S f 中の後寄りの 4 0 0 サンプルのデータと、期間の末尾から 9 2 0 サンプルの後方部データ S b 中の前寄りの 4 0 0 サンプルのデータとを、フェード期間として、前方部データ S f と後方部データ S b をクロスフェードによって接合する。このようにクロスフェードによって接合するのは、接合された部分で圧縮後のデータにノイズが発生するのを防止するためである。

【 0 0 3 8 】

処理単位期間内に無音部分が存在するときには、その無音部分を含めた 1 0 4 0 サンプルのデータを削除して、それぞれ 9 2 0 サンプルの有音部分を同様のクロスフェードによって接合する。無音部分が 1 4 4 0 サンプル未満ではあるが、1 0 4 0 サンプルを超えるときには、4 0 0 サンプルのフェード期間が得られ、接合後のデータとして 1 4 4 0 サンプルのデータが得られるように、有音部分の一部として無音部分の一部を含ませる。

【 0 0 3 9 】

図 3 の処理単位期間 T 1 は、入力データ S i 中に破線で囲んだ有音部分と破線で囲まない無音部分が、ちょうど 3 0 m 秒分（1 4 4 0 サンプル）ずつ存在するとともに、再生開始直後であるため、出力バッファ 2 4 にはデータが全く蓄積されていない場合である。

【 0 0 4 0 】

したがって、処理単位期間 T 1 については、ステップ 3 1, 3 2 および 3 4 によって、図 3 の出力データ S o 中の部分 E 1 に示すように、ちょうど 3 0 m 秒分（1 4 4 0 サンプル）の無音部分が全て削除され、無音部分の前後の有音部分が接合されて、ちょうど 3 0 m 秒分（1 4 4 0 サンプル）の有音部分のデータのみが全て、出力バッファ 2 4 に出力され、書き込まれる。

【0 0 4 1】

ただし、図 3 は入力データ  $S_i$  と出力データ  $S_o$  のタイミング関係を示すものではなく、タイミング的には出力データ  $S_o$  は入力データ  $S_i$  に対して同図の位置より遅れる。

【0 0 4 2】

出力バッファ 2 4 に書き込まれた処理単位期間  $T_1$  の 3 0 m 秒分（1 4 4 0 サンプル）の有音部分のデータは、期間  $T_1$  後の 6 0 m 秒の時間に渡って、出力バッファ 2 4 から読み出され、D/A コンバータ 2 5 によってアナログ音声信号に変換される。

【0 0 4 3】

図 3 の処理単位期間  $T_2$  は、6 0 m 秒分（2 8 8 0 サンプル）のデータが全て有音部分であるとともに、出力バッファ 2 4 にこの 6 0 m 秒分（2 8 8 0 サンプル）の有音部分のデータを全て蓄えるだけの空きがある場合である。

【0 0 4 4】

したがって、処理単位期間  $T_2$  については、ステップ 3 1, 3 5 および 3 3 によって、図 3 の出力データ  $S_o$  中の部分 E 2 に示すように、その 6 0 m 秒分（2 8 8 0 サンプル）の有音部分のデータがそのまま、出力バッファ 2 4 に出力され、書き込まれる。

【0 0 4 5】

この出力バッファ 2 4 に書き込まれた処理単位期間  $T_2$  の 6 0 m 秒分（2 8 8 0 サンプル）の有音部分のデータ中の、前半の 3 0 m 秒分（1 4 4 0 サンプル）のデータが、期間  $T_2$  後の 6 0 m 秒の時間に渡って、出力バッファ 2 4 から読み出され、D/A コンバータ 2 5 によってアナログ音声信号に変換される。

【0 0 4 6】

図 3 の処理単位期間  $T_3$  は、入力データ  $S_i$  中に破線で囲んだ有音部分が 3 0 m 秒分（1 4 4 0 サンプル）未満しか存在せず、破線で囲まない無音部分が 3 0 m 秒分（1 4 4 0 サンプル）を超えて存在するとともに、出力バッファ 2 4 に 3 0 m 秒分（1 4 4 0 サンプル）以上のデータが蓄積されている場合である。

## 【 0 0 4 7 】

したがって、処理単位期間 T 3 については、ステップ 3 1, 3 2 および 3 3 によって、図 3 の出力データ S o 中の部分 E 3 に示すように、その 3 0 m 秒分 ( 1 4 4 0 サンプル ) を超える無音部分が全て削除され、無音部分の前後の有音部分が接合されて、3 0 m 秒分 ( 1 4 4 0 サンプル ) 未満の有音部分のデータのみが全て、出力バッファ 2 4 に出力され、書き込まれる。

## 【 0 0 4 8 】

そして、処理単位期間 T 2 の 6 0 m 秒分 ( 2 8 8 0 サンプル ) の有音部分のデータ中の、出力バッファ 2 4 に蓄積されていた後半の 3 0 m 秒分 ( 1 4 4 0 サンプル ) のデータが、処理単位期間 T 3 後の 6 0 m 秒の時間に渡って、出力バッファ 2 4 から読み出され、D / A コンバータ 2 5 によってアナログ音声信号に変換される。

## 【 0 0 4 9 】

図 3 の処理単位期間 T 4 は、6 0 m 秒分 ( 2 8 8 0 サンプル ) のデータが全て有音部分であるとともに、出力バッファ 2 4 にこの 6 0 m 秒分 ( 2 8 8 0 サンプル ) の有音部分のデータを全て蓄えるだけの空きがある場合である。

## 【 0 0 5 0 】

したがって、処理単位期間 T 4 については、ステップ 3 1, 3 5 および 3 3 によって、図 3 の出力データ S o 中の部分 E 4 に示すように、その 6 0 m 秒分 ( 2 8 8 0 サンプル ) の有音部分のデータがそのまま、出力バッファ 2 4 に出力され、書き込まれる。

## 【 0 0 5 1 】

そして、出力バッファ 2 4 に蓄積されていた処理単位期間 T 3 の 3 0 m 秒分 ( 1 4 4 0 サンプル ) 未満の有音部分のデータと、出力バッファ 2 4 に書き込まれた処理単位期間 T 4 の 6 0 m 秒分 ( 2 8 8 0 サンプル ) の有音部分のデータ中の前半の一部のデータとを合わせた、3 0 m 秒分 ( 1 4 4 0 サンプル ) のデータが、期間 T 4 後の 6 0 m 秒の時間に渡って、出力バッファ 2 4 から読み出され、D / A コンバータ 2 5 によってアナログ音声信号に変換される。



## 【0052】

図3の処理単位期間T5は、60m秒分（2880サンプル）のデータが全て有音部分であるとともに、出力バッファ24にこの60m秒分（2880サンプル）の有音部分のデータを全て蓄えるだけの空きがない場合である。

## 【0053】

したがって、処理単位期間T5については、ステップ31、35および36によって、図3の出力データS<sub>o</sub>中の部分E5に示し、かつ図7に示したように、60m秒分（2880サンプル）の有音部分のデータが、30m秒分（1440サンプル）に圧縮されて、出力バッファ24に出力され、書き込まれる。

## 【0054】

そして、処理単位期間T4の60m秒分（2880サンプル）の有音部分のデータ中の前半の残りの部分のデータと、後半の一部のデータとを合わせた、30m秒分（1440サンプル）のデータが、処理単位期間T5後の60m秒の時間に渡って、出力バッファ24から読み出され、D/Aコンバータ25によってアナログ音声信号に変換される。

## 【0055】

さらに、処理単位期間T4の60m秒分（2880サンプル）の有音部分のデータ中の後半の残りの部分のデータと、処理単位期間T5の30m秒分（1440サンプル）に圧縮された有音部分のデータ中の一部のデータとを合わせた、30m秒分（1440サンプル）のデータが、処理単位期間T5後の処理単位期間T6後の60m秒の時間に渡って、出力バッファ24から読み出され、D/Aコンバータ25によってアナログ音声信号に変換される。

## 【0056】

以上のように、図3の例では、処理単位期間T1～T4については、入力データS<sub>i</sub>中の無音部分のみが全て削除され、有音部分のデータのみが全て出力バッファ24に書き込まれるとともに、出力バッファ24からは有音部分のデータのみが全て、60m秒で1440サンプルの速度で読み出されるので、出力倍率は1倍となり、有音部分のみが全て定常速度で出力される。

【 0 0 5 7 】

処理単位期間 T 5 については、6 0 m 秒分（2 8 8 0 サンプル）の有音部分のデータが、3 0 m 秒分（1 4 4 0 サンプル）に圧縮されて出力バッファ 2 4 に書き込まれ、6 0 m 秒の時間で出力バッファ 2 4 から読み出されるので、例外的に出力倍率は 2 倍となり、速聴きとなる。

【 0 0 5 8 】

なお、図 6 のステップ 3 4 で、有音部分の全てに無音部分を一部付加して、3 0 m 秒分（1 4 4 0 サンプル）のデータを出力バッファ 2 4 に書き込む場合については、図 3 には示してしないが、例えば、図 3 の処理単位期間 T 2 では、出力バッファ 2 4 に 3 0 m 秒分（1 4 4 0 サンプル）以上のデータが蓄積されていないので、期間 T 2 が図 3 とは異なり、有音部分が 3 0 m 秒分（1 4 4 0 サンプル）未満しか存在せず、無音部分が 3 0 m 秒分（1 4 4 0 サンプル）を超えて存在する場合には、その 3 0 m 秒分（1 4 4 0 サンプル）未満の有音部分の全てに、期間 T 2 の先頭に近い部分から、3 0 m 秒分（1 4 4 0 サンプル）を超える無音部分の一部が付加されて、期間 T 2 内の 3 0 m 秒分（1 4 4 0 サンプル）のデータが、出力バッファ 2 4 に出力され、書き込まれる。

【 0 0 5 9 】

以上のように、上述した例によれば、出力バッファ 2 4 を効果的に用いることによって、必要最小限のメモリを用いるだけで、音声が続切れず、内容を理解でき、聴きやすい状態で、定常速度より速い速度で再生された音声信号を定常速度に変換して出力することができる。

【 0 0 6 0 】

〔データ圧縮の際の位置調整…図 8、図 9〕

図 6 のステップ 3 6 で、図 3 の部分 E 5 および図 7 の下段に示したようにクロスフェードによって前方部データ S f と後方部データ S b を接合する場合、図 8 (A) に示すように、フェード期間において前方部データ S f と後方部データ S b の位相が一致しているときには、接合後のデータは接合部分でうなりを生じない。

【0061】

しかし、図8（B）に示すように、フェード期間において前方部データ S f と後方部データ S b の位相が大きくずれているときには、接合後のデータは接合部分でうなりを生じ、音質が劣化する。

【0062】

そこで、クロスフェードによって前方部データ S f と後方部データ S b を接合するに当たっては、以下のように前方部データ S f と後方部データ S b の位置関係を調整する。

【0063】

すなわち、図9に示すように、後方部データ S b を、前方部データ S f に対して、所定サンプル数の時間内で、1 サンプルまたは数サンプルの時間ずつ、タイミング的に移動させて、それぞれの移動位置において、400 サンプルのフェード期間内の前方部データ S f のデータ値と後方部データ S b のデータ値との差分の絶対値の、接合後のデータとして示す波形の塗り潰した部分の面積で表されるフェード期間全体に渡る積分値（和）を算出する。

【0064】

そして、この積分値が最小となる移動位置を、前方部データ S f と後方部データ S b が最適な位相関係になる位置として検出し、その移動位置において、上述したようにクロスフェードによって前方部データ S f と後方部データ S b を接合する。

【0065】

前方部データ S f および後方部データ S b のフェード期間におけるデータを、それぞれ F [n] および B [n] とすると、接合後のデータの接合部分（フェード期間）におけるデータ X [n] は、

$$X[n] = ((400 - n) * F[n] + n * B[n]) / 400 \dots (1)$$

で表される。ただし、n = 0, 1, 2...399、である。

【0066】

この例によれば、接合後のデータは接合部分でうなりや違和感の少ないものとなる。

## 【0067】

〔無音部分削除の際のノイズ発生防止…図10、図11〕

図6のステップ34で図3の部分E1に示したように、または図6のステップ33で図3の部分E3に示したように、入力データ $S_i$ 中の無音部分を削除して、前後の有音部分を接合する場合、無音部分として図10の上段に示すデータ部分 $S_p$ を削除したときには、部分 $S_p$ は無音部分でレベルが低いものの、図10の下段左側に示すように、接合後の出力データ $S_o$ は接合点の前後で位相が大きく変化するため、接合点でボツ音と呼ばれるノイズが発生する。また、無音部分として図10の上段に示すデータ部分 $S_q$ を削除したときには、図10の下段右側に示すように、同様に接合後の出力データ $S_o$ には接合点でボツ音と呼ばれるノイズが発生する。

## 【0068】

そこで、削除した無音部分の前後の有音部分を接合するに当たっては、接合点の前後のそれぞれ所定サンプル数 $k$ の有音部分につき、フェードアウト処理およびフェードイン処理を行う。

## 【0069】

具体的に、図10下段の左側および右側に示すように、フェードアウト処理前の接合部前半のデータを $C1[i]$ 、フェードイン処理前の接合部後半のデータを $D1[i]$ とし、図11の左側および右側に示すように、フェードアウト処理後の接合部前半のデータを $C2[i]$ 、フェードイン処理後の接合部後半のデータを $D2[i]$ とすると、

$$C2[i] = (k - i) * C1[i] / k \quad \dots (2)$$

$$D2[i] = i * D1[i] / k \quad \dots (3)$$

とする。ただし、 $i = 0, 1, 2 \dots (k - 1)$ 、である。

## 【0070】

このようにフェードアウト処理およびフェードイン処理をして、前後の有音部分を接合することによって、図11に示すように、接合後の出力データ $S_o$ は接合点でボツ音と呼ばれるノイズが発生しなくなる。

## 【 0 0 7 1 】

〔音程に応じて処理単位時間を変化させる場合…図 1 2、図 1 3〕

図 1 2 は、音声処理部 9 の他の例を示し、音声信号の主成分の音程に応じて速度変換処理の処理単位時間を変化させる場合である。音程に対して処理単位時間を直線的に変化させることも可能であるが、以下の例は、音程を 3 段階に分けて処理単位時間を 3 段階に変化させる場合である。

## 【 0 0 7 2 】

この例では、音程検出部 2 7 において、倍速デコーダ 7 の出力の音声信号の主成分（レベルが最大の周波数成分）の音程が検出される。音程検出部 2 7 での音程検出方法には、FFT（高速フーリエ変換）などの公知の方法を用いることができる。

## 【 0 0 7 3 】

そして、この例では、処理単位期間設定部 2 1 は、音程検出部 2 7 の検出結果に基づいて、図 1 3 に示す音程対応処理単位期間設定処理ルーチンによって、処理単位時間を決定し、処理単位期間の先頭ないし末尾のタイミングを決定する。

## 【 0 0 7 4 】

すなわち、この音程対応処理単位期間設定処理ルーチン 4 0 では、まずステップ 4 1 において、音程検出部 2 7 の検出結果の音程が 3 0 0 H z 以上であるか否かを判断し、3 0 0 H z 以上であるときには、ステップ 4 1 からステップ 4 2 に進んで、処理単位時間を最短の 4 0 m 秒として、当該の処理単位期間の先頭（直前の処理単位期間の末尾）から 4 0 m 秒後を当該の処理単位期間の末尾（直後の処理単位期間の先頭）とする。

## 【 0 0 7 5 】

検出結果の音程が 3 0 0 H z 未満であるときには、ステップ 4 1 からステップ 4 3 に進んで、検出結果の音程が 1 0 0 H z 以下であるか否かを判断し、1 0 0 H z 以下であるときには、ステップ 4 3 からステップ 4 4 に進んで、処理単位時間を最長の 6 0 m 秒として、当該の処理単位期間の先頭から 6 0 m 秒後を当該の処理単位期間の末尾とする。

## 【0076】

検出結果の音程が100Hz以下でもないときには、すなわち100Hzを超え、かつ300Hz未満であるときには、ステップ43からステップ45に進んで、処理単位時間を中間の50m秒として、当該の処理単位期間の先頭から50m秒後を当該の処理単位期間の末尾とする。

## 【0077】

図12の速度変換処理部23が行う速度変換処理は、図3の処理単位期間T1, T2, T3…が、60m秒という固定の時間ではなく、40m秒、50m秒または60m秒というように変化し、図6の速度変換処理ルーチン30における倍率相当分が、処理単位時間の変化に応じて、2倍速再生の場合には20m秒、25m秒または30m秒というように変化する点を除いて、上述した例と同じである。

## 【0078】

この例によれば、音声信号の主成分の音程が高いときには速度変換処理の処理単位時間が短くなり、主成分の音程が低いときには速度変換処理の処理単位時間が長くなるので、音声信号の周波数による歪みや揺れの少ない出力音声信号を得ることができる。

## 【0079】

〔レベルが低い点を処理単位期間の区切れ目とする場合…図14～図16〕

図14は、音声処理部9のさらに他の例を示し、無音部分と有音部分の境目、または有音部分中の相対的にレベルが低い部分を、速度変換処理の処理単位期間の区切れ目とする場合である。

## 【0080】

この例では、処理単位期間設定部21は、倍速デコーダ7の出力の音声信号、および有音無音判定部22の判定結果を取り込んで、図15に示す音声レベル対応処理単位期間設定処理ルーチンによって、処理単位期間の先頭ないし末尾のタイミングを決定する。

## 【0081】

すなわち、この音声レベル対応処理単位期間設定処理ルーチン50では、まず

ステップ 5 1 において、有音無音判定部 2 2 の判定結果から、当該の処理単位期間の先頭から 3 0 m 秒以上 1 0 0 m 秒以下の時間範囲内に、無音部分と有音部分の切り替わり時点（無音部分から有音部分に切り替わる時点、または有音部分から無音部分に切り替わる時点）が存在するか否かを判断する。

【 0 0 8 2 】

そして、その時間範囲内に無音部分と有音部分の切り替わり時点が存在するときには、ステップ 5 1 からステップ 5 2 に進んで、その無音部分と有音部分の切り替わり時点を当該の処理単位期間の末尾とする。その時間範囲内に無音部分と有音部分の切り替わり時点が 2 点以上存在するときには、最初の時点を当該の処理単位期間の末尾とする。

【 0 0 8 3 】

一方、その時間範囲内に無音部分と有音部分の切り替わり時点が存在しないときには、ステップ 5 1 からステップ 5 3 に進んで、その時間範囲内で音声平均レベルが最小となる時点当該の処理単位期間の末尾とする。音声平均レベルは、図 4 に示したような判定単位期間内の各サンプルのデータ値の絶対値の平均値であり、その平均値が最小となる判定単位期間の先頭または末尾を当該の処理単位期間の末尾とする。

【 0 0 8 4 】

したがって、この例では、処理単位時間は、最短で 3 0 m 秒、最長で 1 0 0 m 秒である。

【 0 0 8 5 】

図 1 6 に「固定長の場合」として示すケースは、図 2 の例のように処理単位期間  $T_1$ 、 $T_2$ 、 $T_3 \dots$  が 6 0 m 秒という固定の時間にされる場合であり、図 1 6 に「可変長の場合」として示すケースは、この図 1 4 の例のように処理単位期間  $T_{11}$ 、 $T_{12}$ 、 $T_{13} \dots$  の先頭ないし末尾のタイミングが決定され、時間が変えられる場合である。

【 0 0 8 6 】

すなわち、この例では、時間軸上で 0 m 秒として示す最初の処理単位期間  $T_1$  の先頭の時点から 3 0 m 秒以上 1 0 0 m 秒以下の時間範囲内に、入力データ  $S$

i が有音部分から無音部分に切り替わる時点  $t_a$  が存在するので、その時点  $t_a$  が、処理単位期間 T 1 1 の末尾、すなわち次の処理単位期間 T 1 2 の先頭となる。

【0087】

さらに、処理単位期間 T 1 2 の先頭の時点  $t_a$  から 30 m 秒以上 100 m 秒以下の時間範囲内に、入力データ  $S_i$  が無音部分から有音部分に切り替わる時点  $t_b$  が存在するので、その時点  $t_b$  が、処理単位期間 T 1 2 の末尾、すなわち次の処理単位期間 T 1 3 の先頭となる。

【0088】

さらに、処理単位期間 T 1 3 の先頭の時点  $t_b$  から 30 m 秒以上 100 m 秒以下の時間範囲内には無音部分と有音部分の切り替わり時点が存在しないので、その時間範囲内で入力データ  $S_i$  の音声平均レベルが最小となる時点  $t_c$  が、処理単位期間 T 1 3 の末尾、すなわち次の処理単位期間 T 1 4 の先頭となる。

【0089】

この例によれば、無音部分と有音部分の境目、または有音部分中の相対的にレベルが低い部分が、速度変換処理の処理単位期間の区切れ目となるので、音声信号の切り貼りによる音質の劣化を大幅に低減することができる。

【0090】

〔他の実施形態〕

図 1 の実施形態は、記録媒体 1 がディスク状の記録媒体である場合であるが、この発明は、記録媒体（記憶媒体）が磁気テープや半導体メモリなどである場合にも、同様に適用することができる。半導体メモリの場合には、再生ヘッドではなく読み出し回路によってメモリから、これに書き込まれた音声信号を読み出すことは言うまでもない。この場合には、「記録」とは書き込みであり、「再生」とは読み出しである。

【0091】

また、この発明は、記録媒体から、これに記録された音声信号のみを定常速度より速い速度で再生する場合または装置にも、同様に適用することができる。



【 0 0 9 2 】

【発明の効果】

上述したように、この発明によれば、出力バッファを効果的に用いることによって、必要最小限のメモリを用いるだけで、音声が続切れず、内容を理解でき、聴きやすい状態で、定常速度より速い速度で再生された音声信号を定常速度に変換して出力することができる。

【図面の簡単な説明】

【図 1】

この発明の再生装置の一実施形態を示す図である。

【図 2】

音声処理部の一例を示す図である。

【図 3】

速度変換処理の一例を示す図である。

【図 4】

有音部分と無音部分の識別判定の一例を示す図である。

【図 5】

出力バッファの入出力ポインタの動きを示す図である。

【図 6】

速度変換処理ルーチンの一例を示す図である。

【図 7】

データ圧縮の一例を示す図である。

【図 8】

データ圧縮の際に位置調整を行わない場合を示す図である。

【図 9】

データ圧縮の際に位置調整を行う場合の調整方法を示す図である。

【図 1 0】

無音部分削除の際のノイズの発生を示す図である。

【図 1 1】

無音部分削除の際のノイズ発生防止方法を示す図である。

【図 1 2】

音声処理部の他の例を示す図である。

【図 1 3】

音程対応処理単位期間設定処理ルーチンの一例を示す図である。

【図 1 4】

音声処理部のさらに他の例を示す図である。

【図 1 5】

音声レベル対応処理単位期間設定処理ルーチンの一例を示す図である。

【図 1 6】

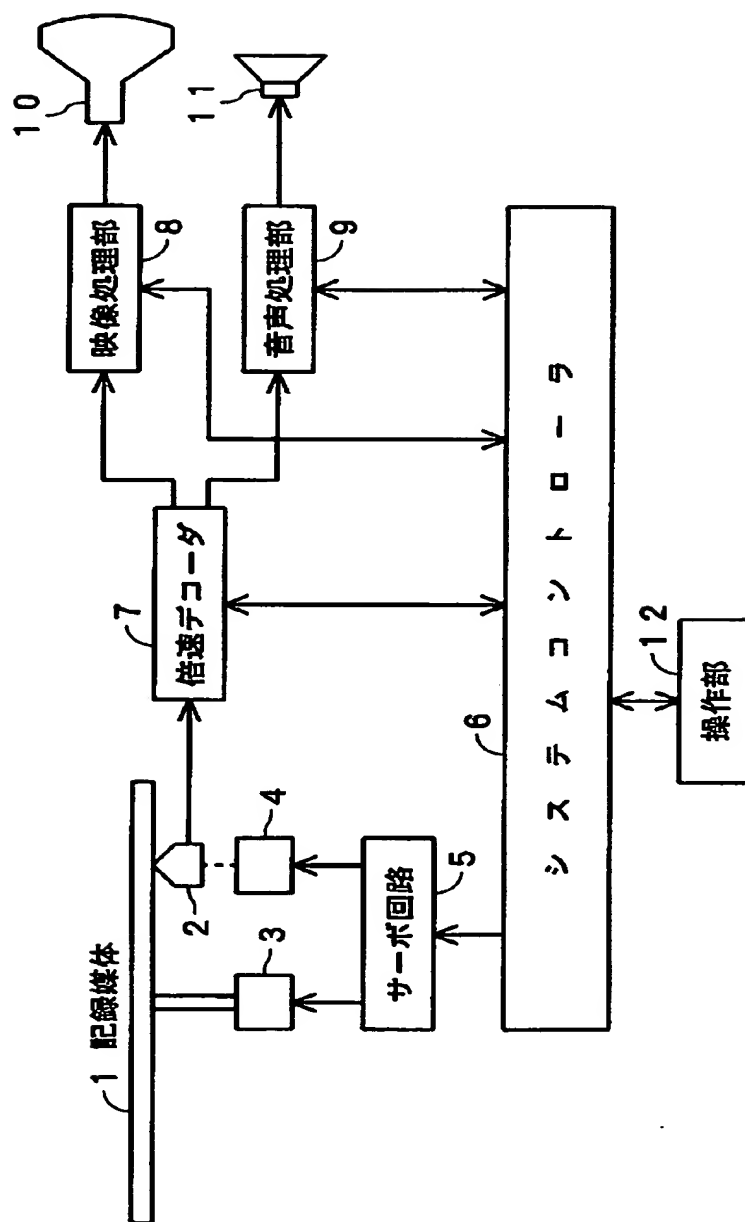
音声レベルが低い点を速度変換処理の処理単位期間の区切れ目とする場合の例を示す図である。

【符号の説明】

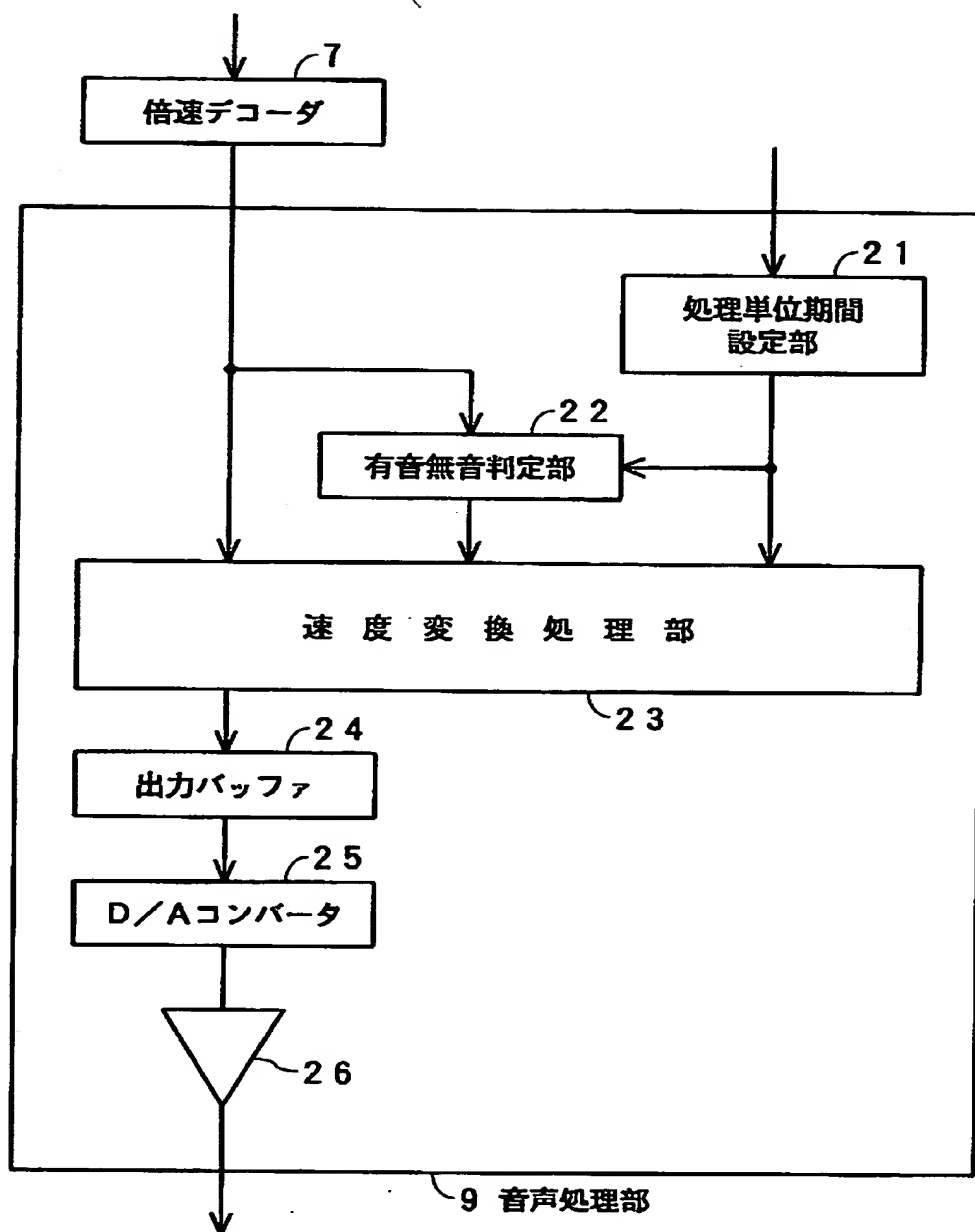
1 …記録媒体、2 …再生ヘッド、7 …倍速デコーダ、8 …映像処理部、9 …音声処理部、21 …処理単位期間設定部、22 …有音無音判定部、23 …速度変換処理部、24 …出力バッファ、25 …D/Aコンバータ、26 …音声増幅器、27 …音程検出部、30 …速度変換処理ルーチン、40 …音程対応処理単位期間設定処理ルーチン、50 …音声レベル対応処理単位期間設定処理ルーチン

【書類名】 図面

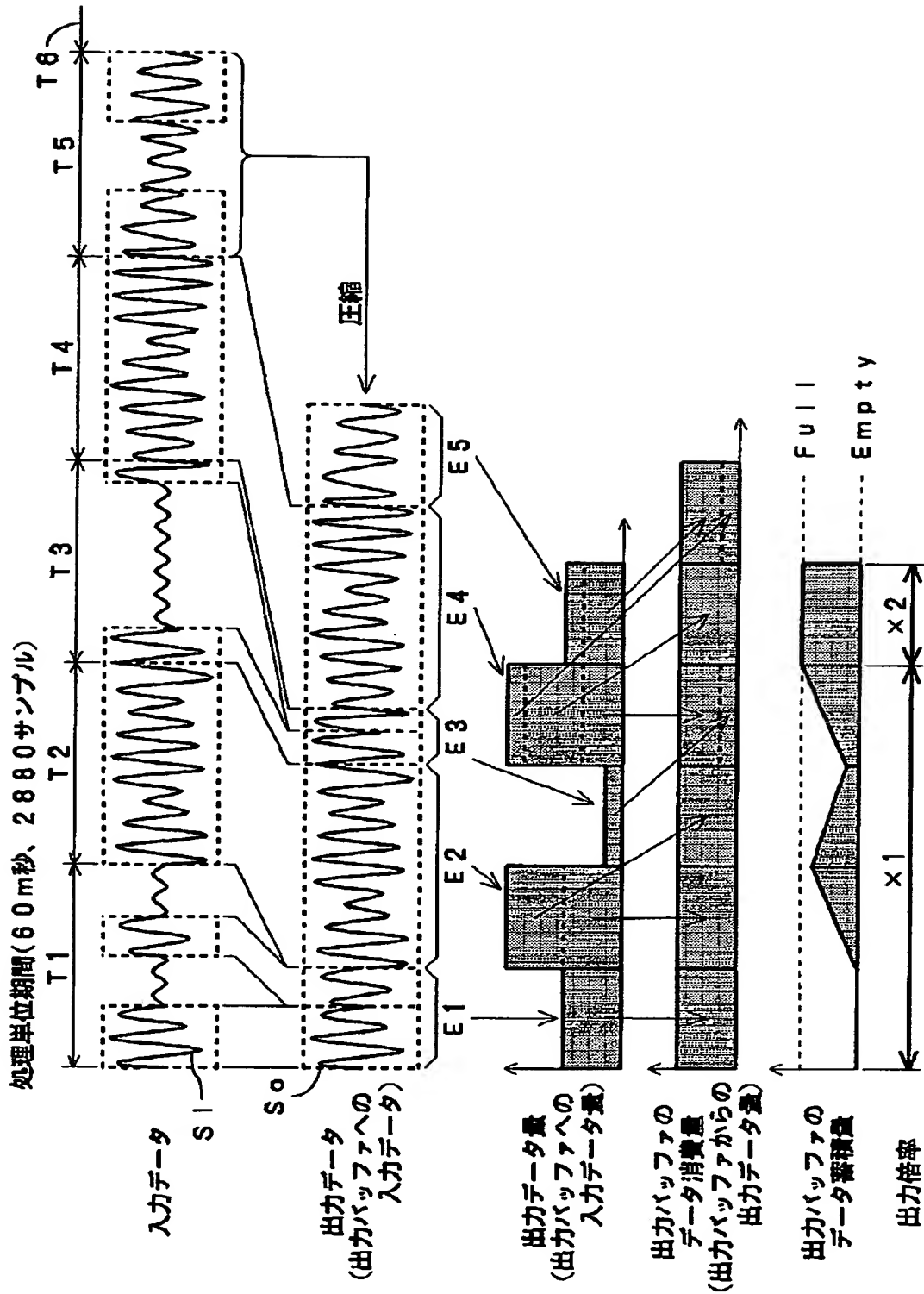
【図 1】



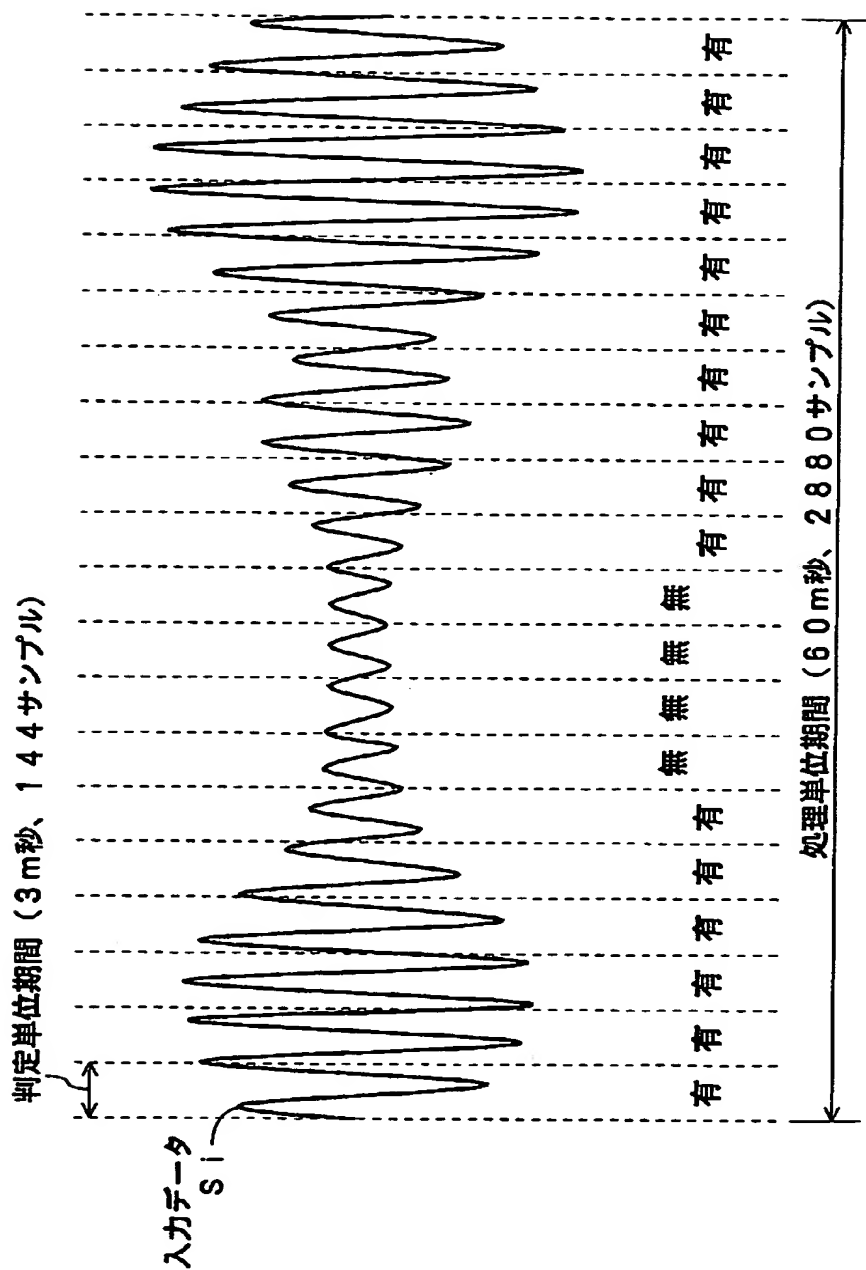
【図 2】



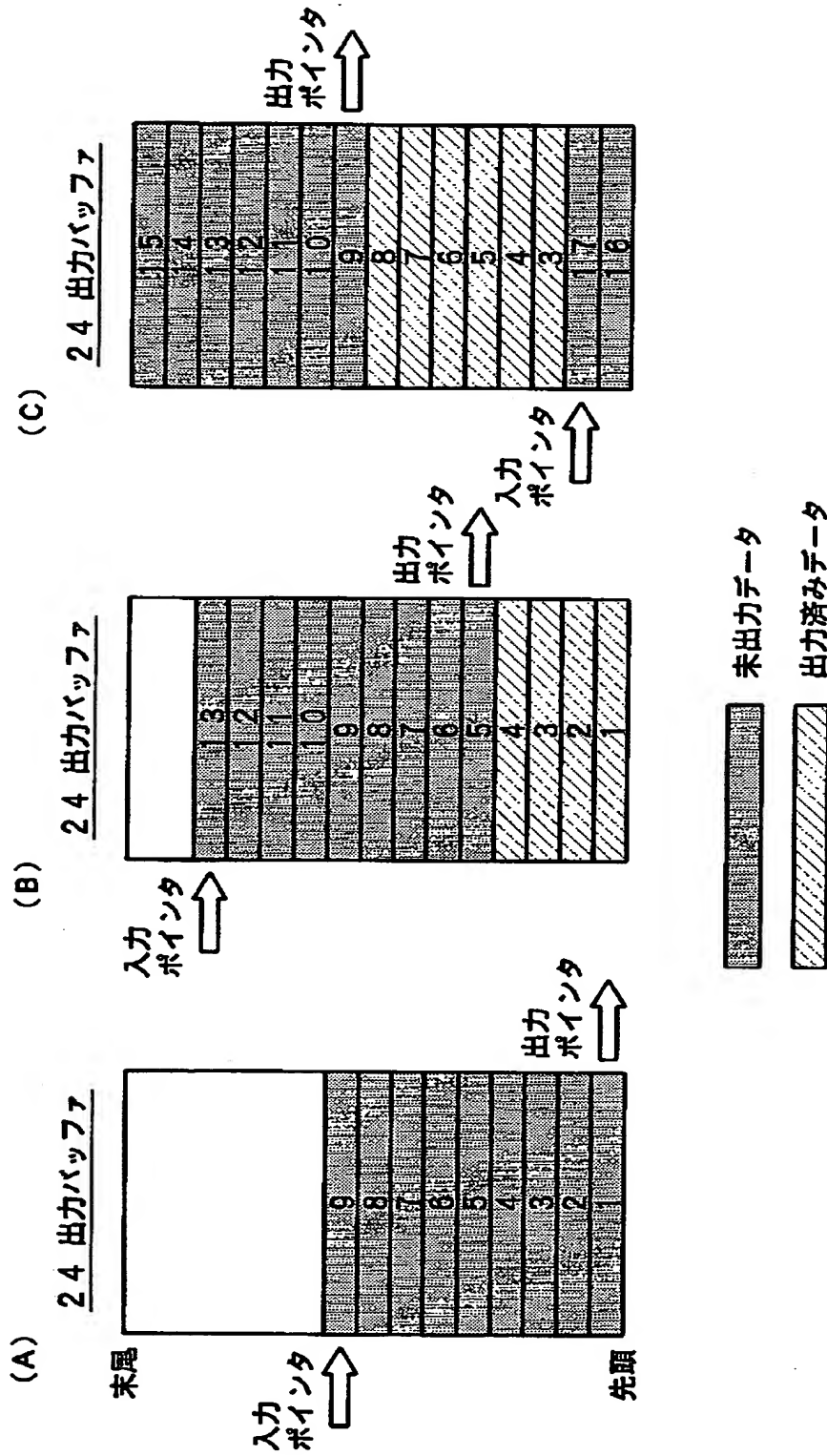
【図 3】



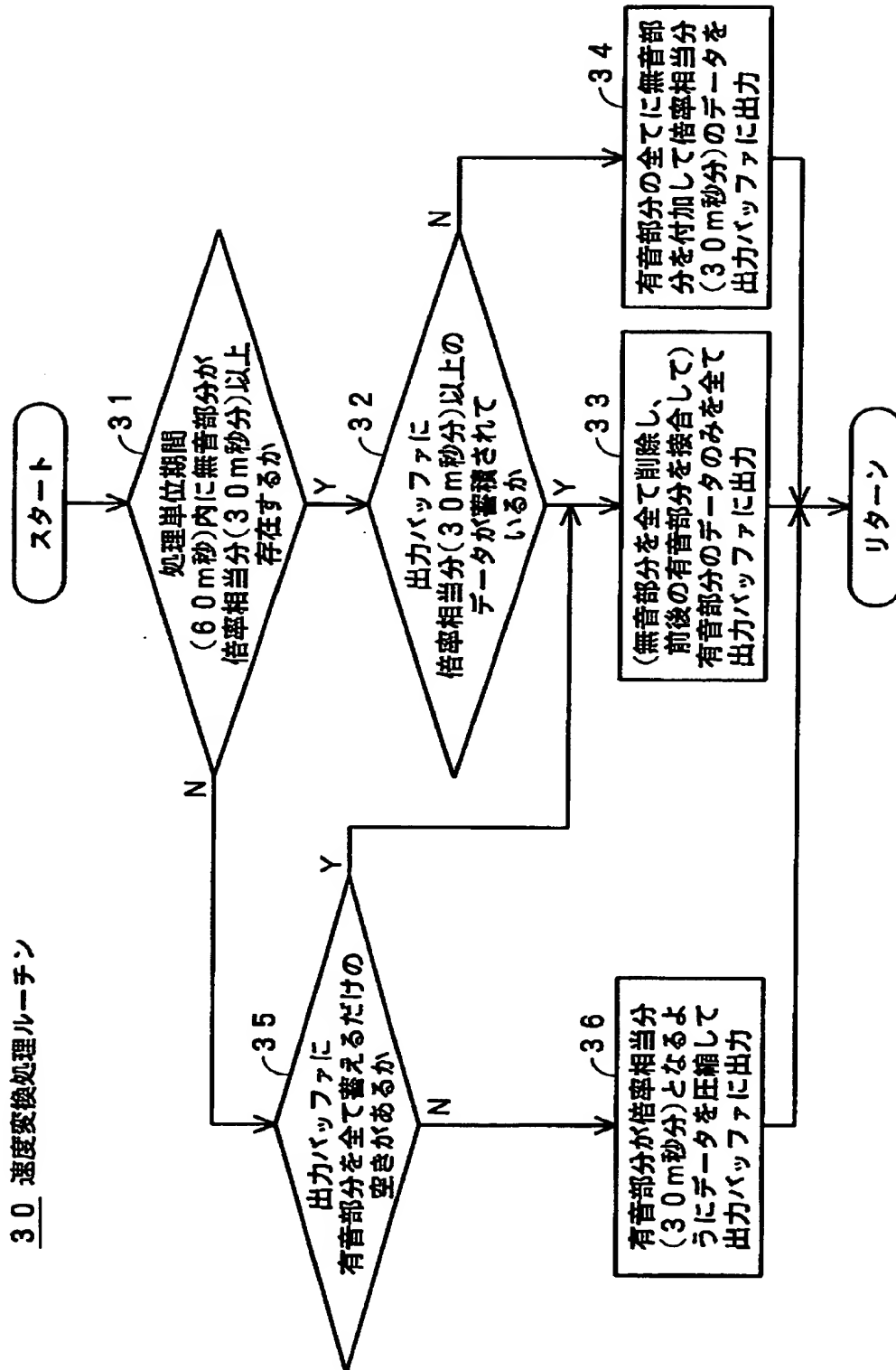
【図 4】



【図 5】

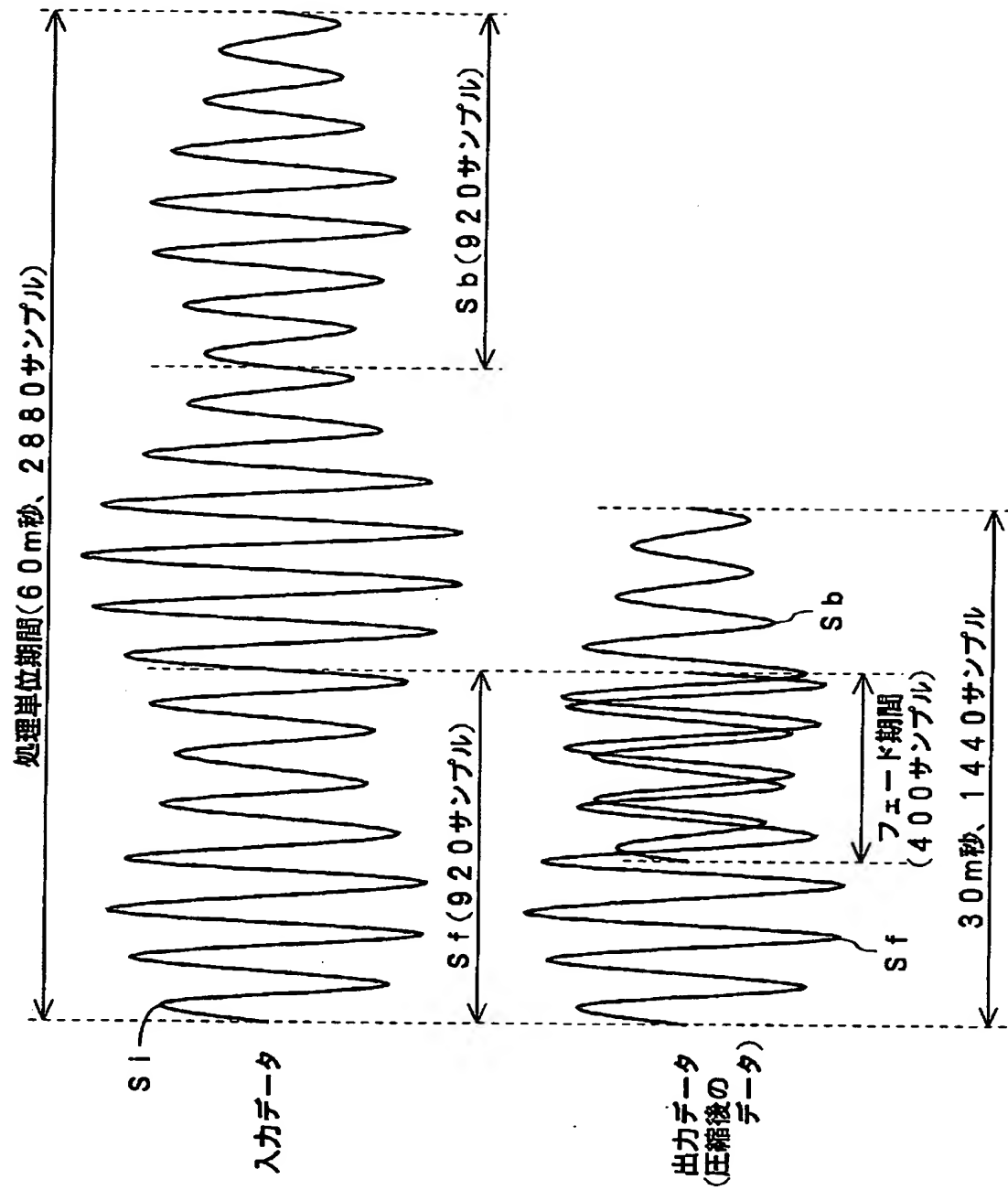


【図 6】

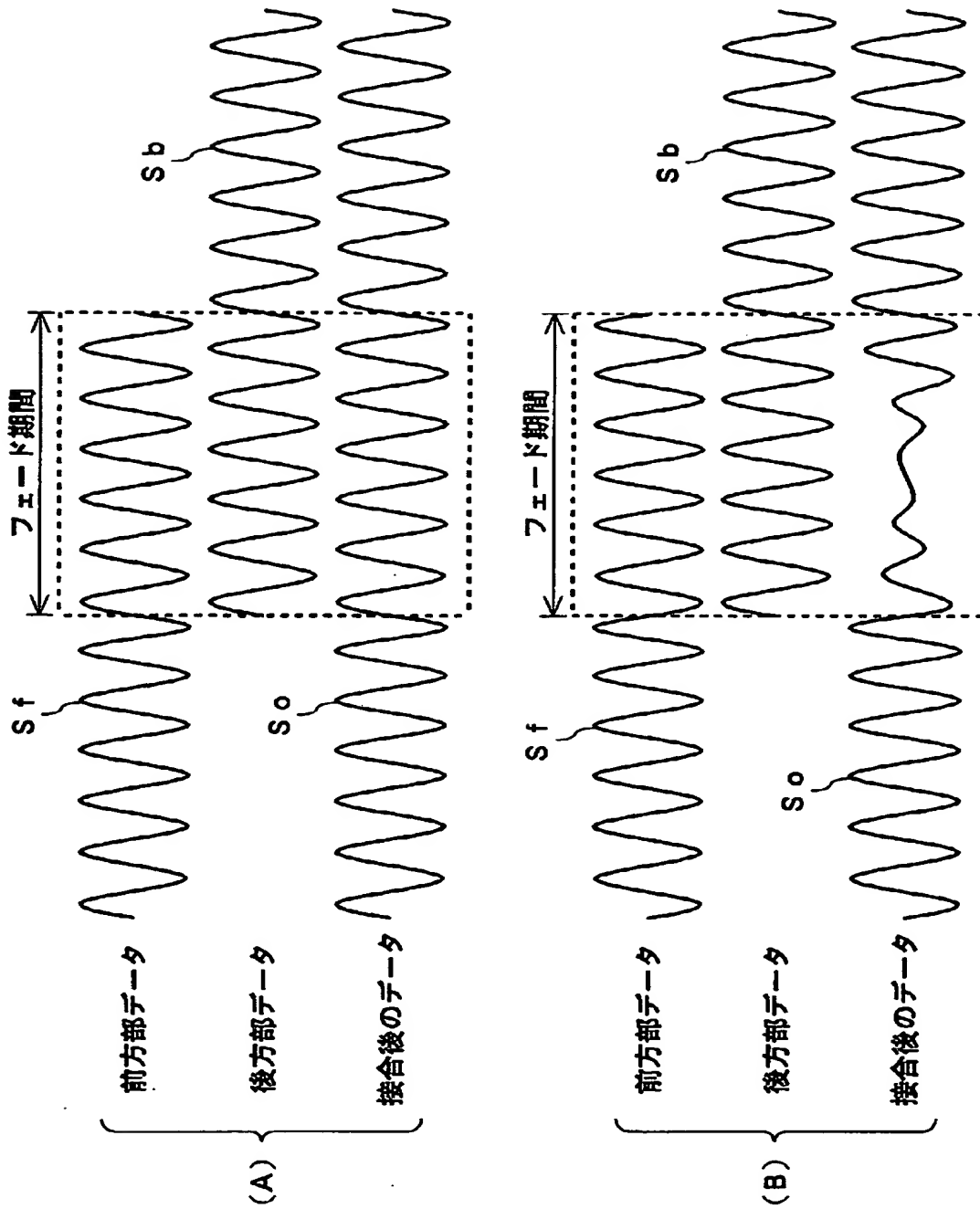




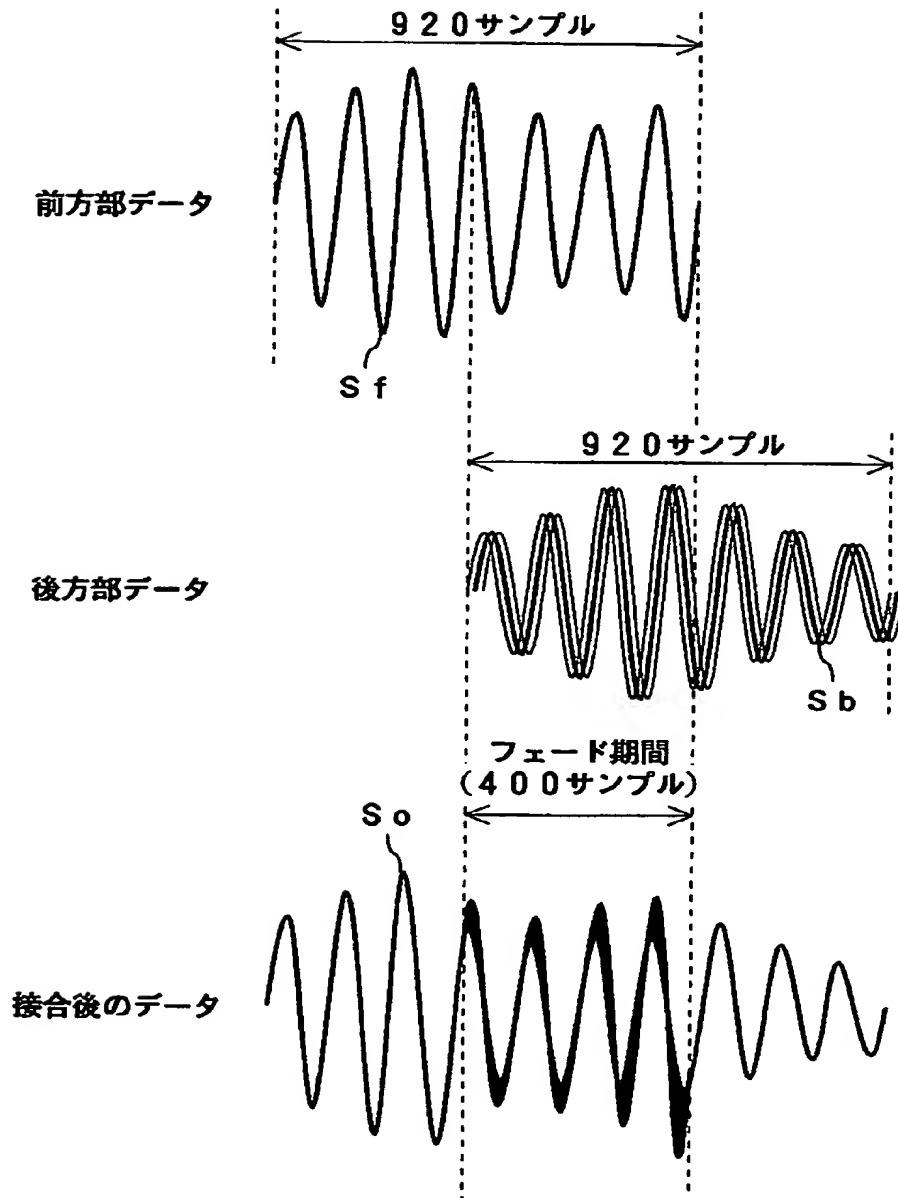
【図 7】



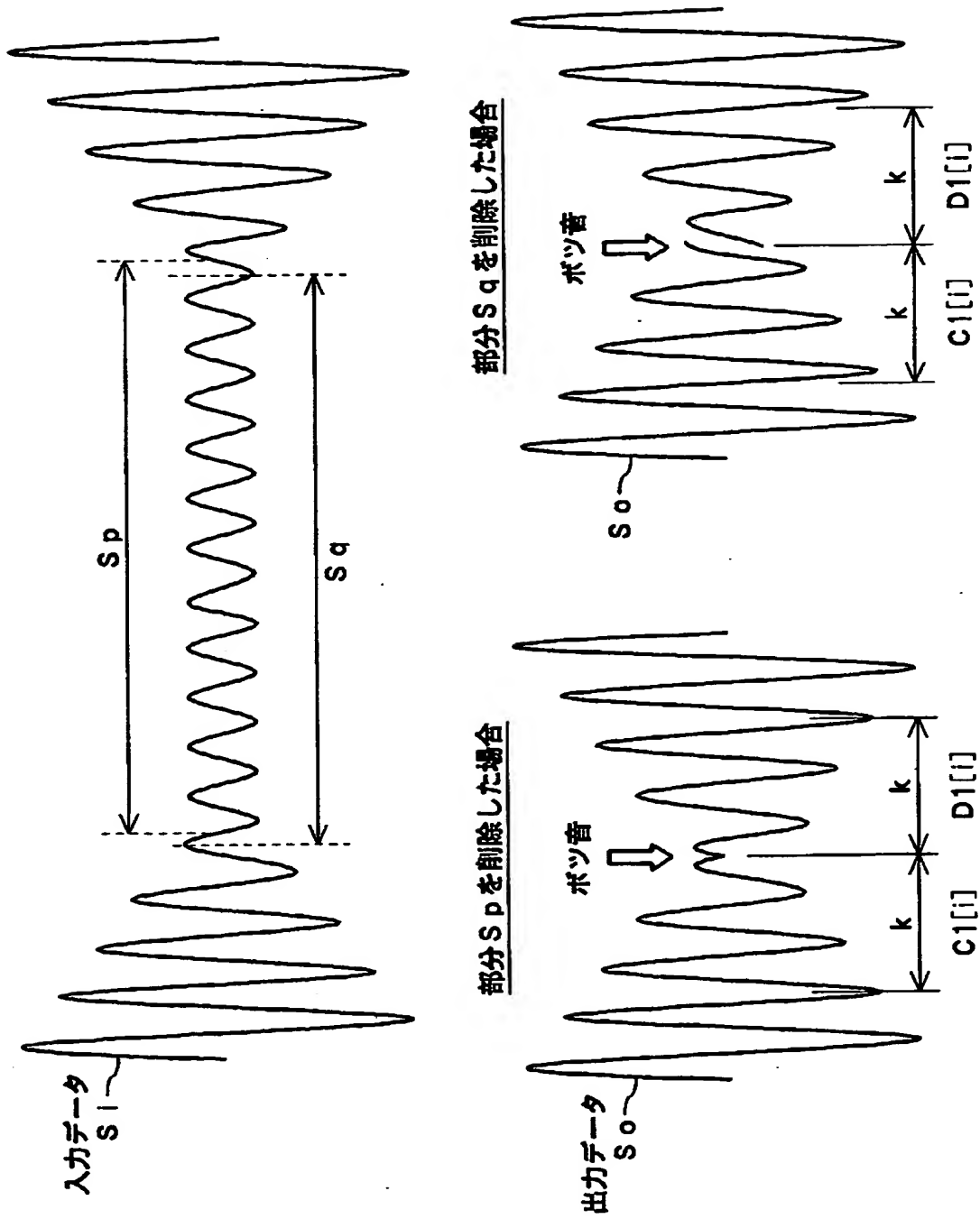
【図 8】



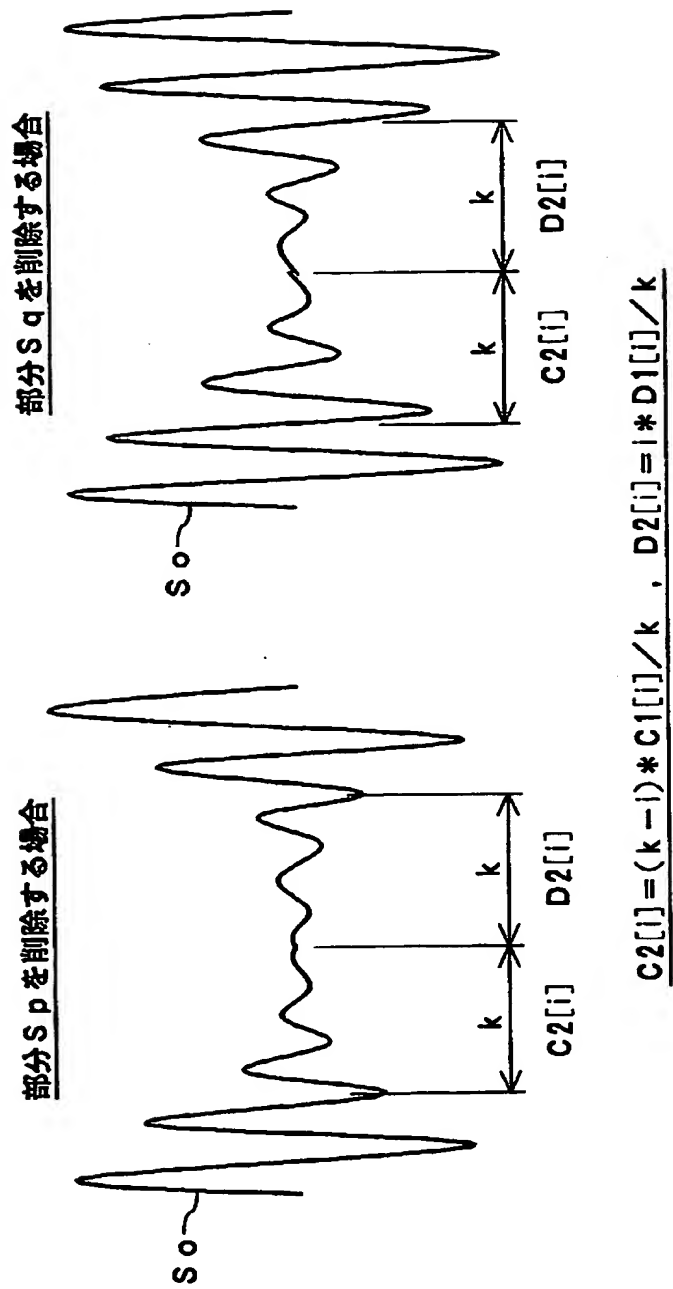
【図 9】



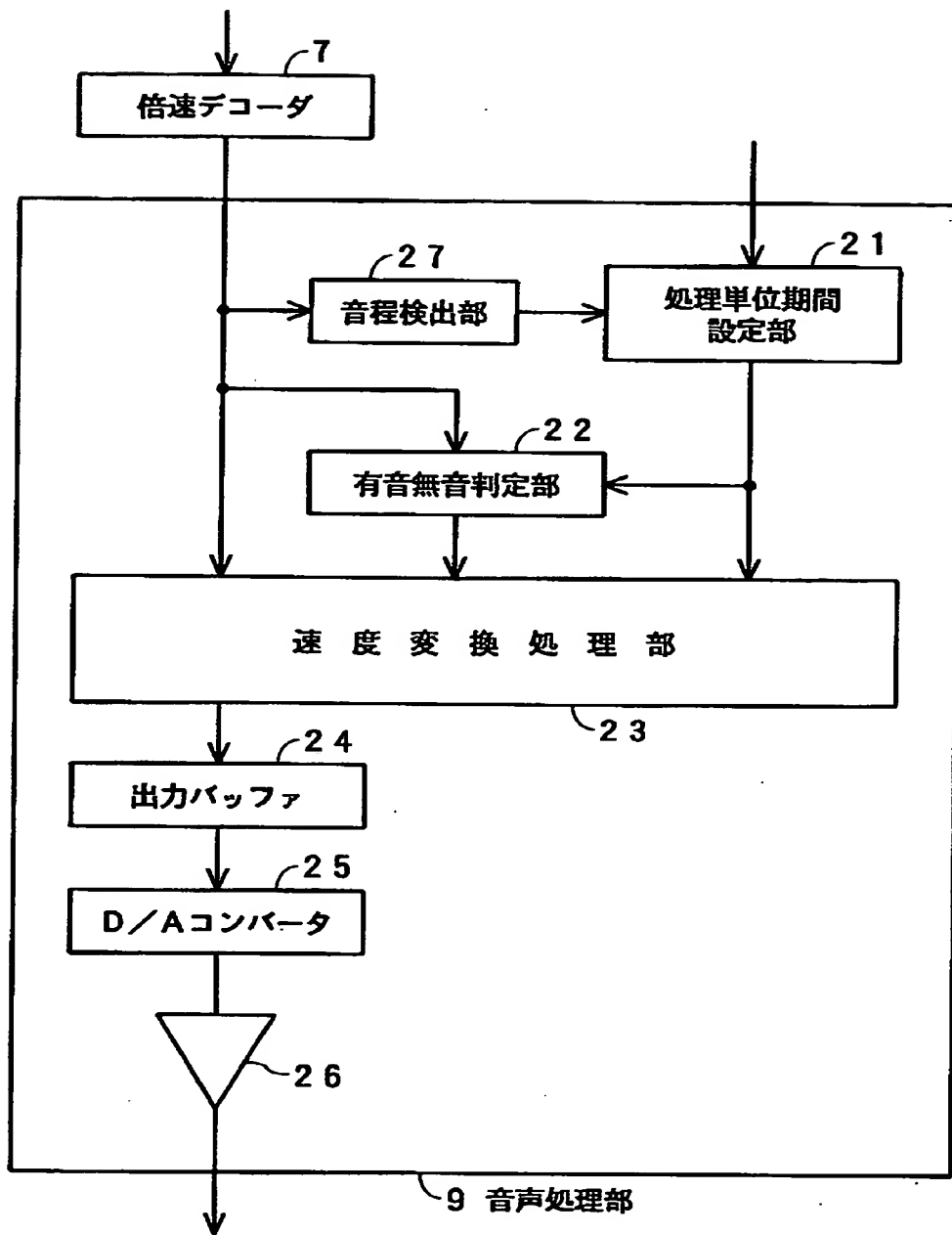
【図 1 0】



【図 1 1】

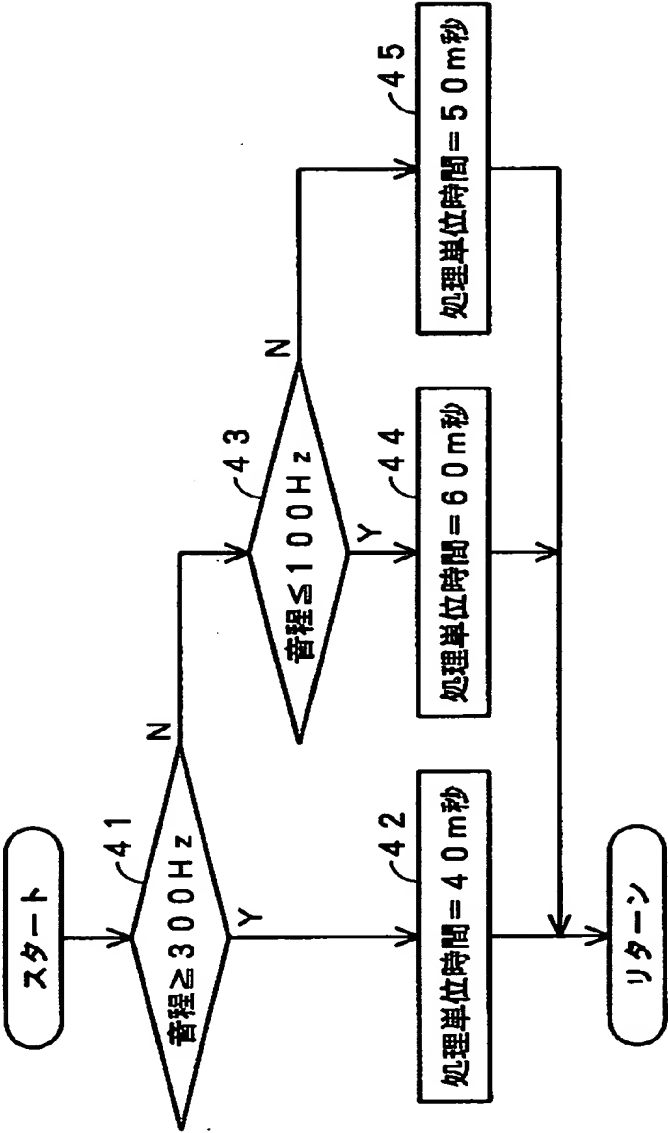


【図 1 2】

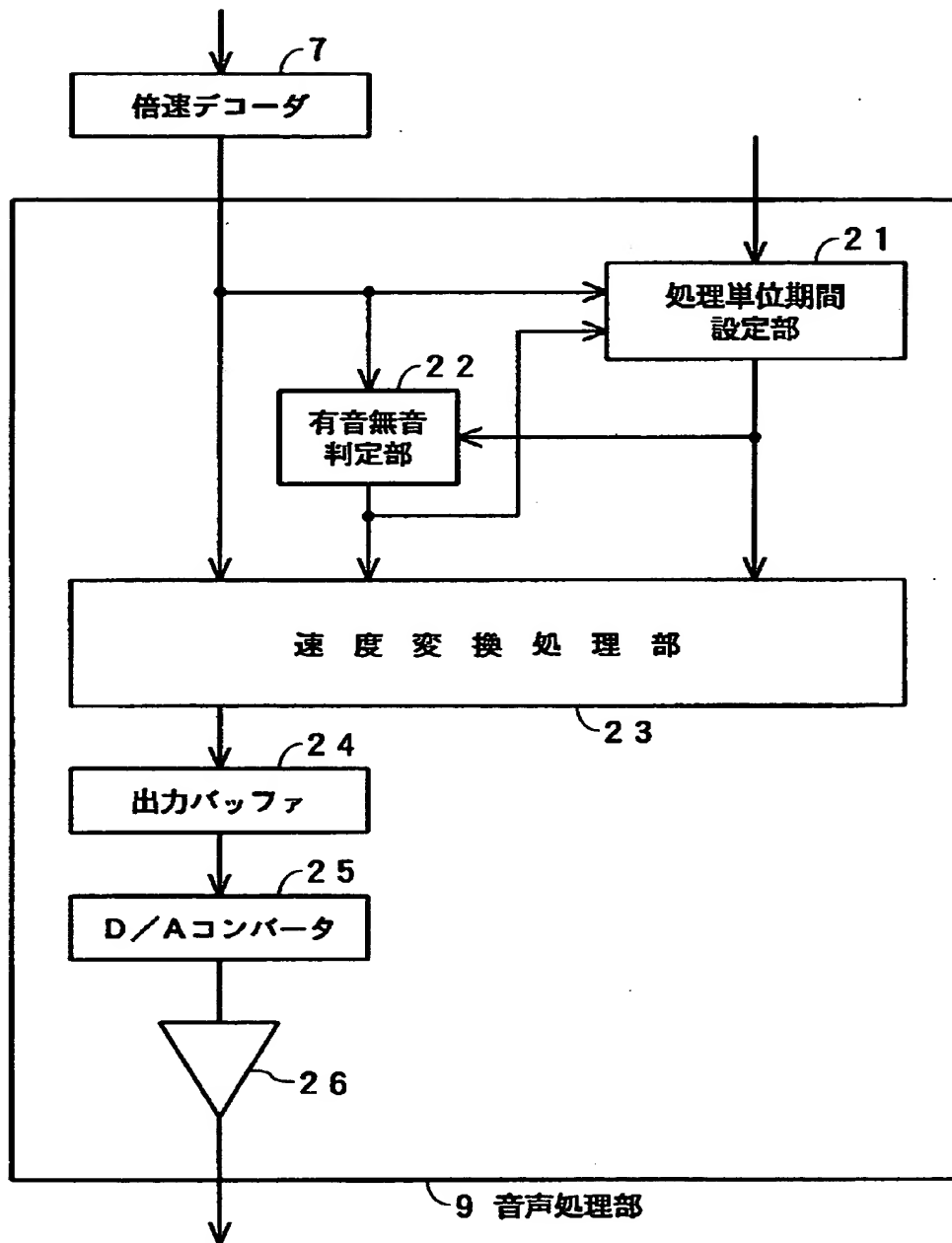


【図 1 3】

4 0 音程対応処理単位期間設定処理ルーチン



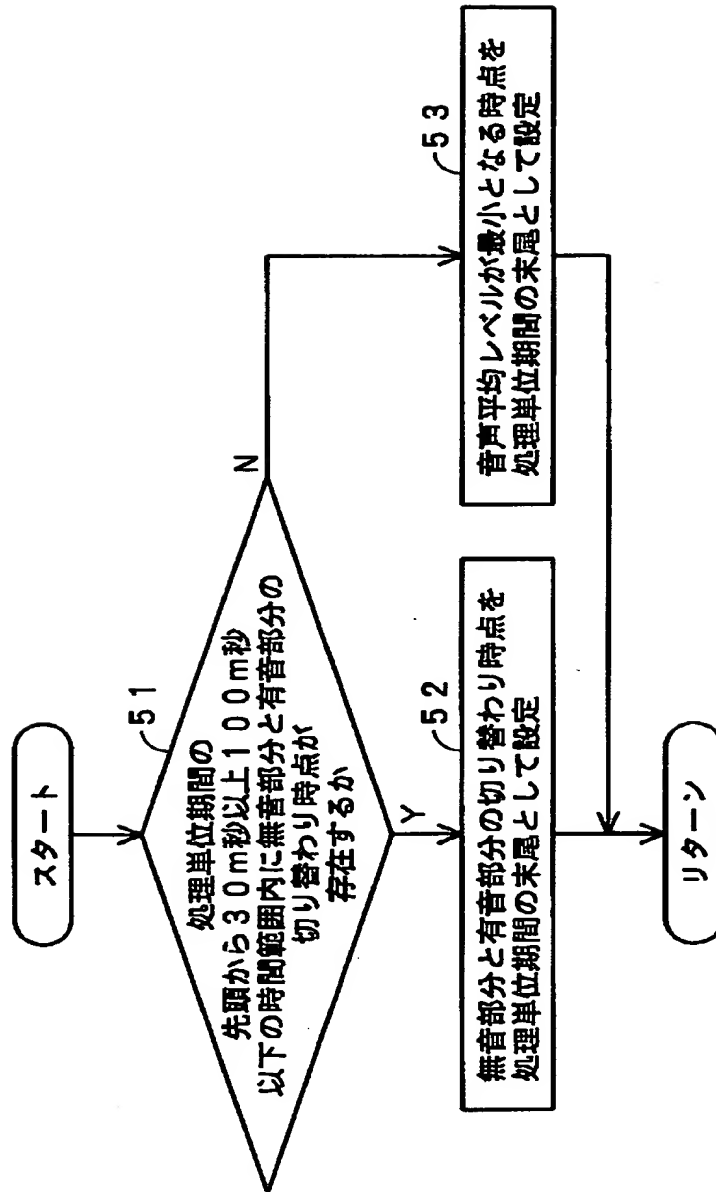
【図 1 4】



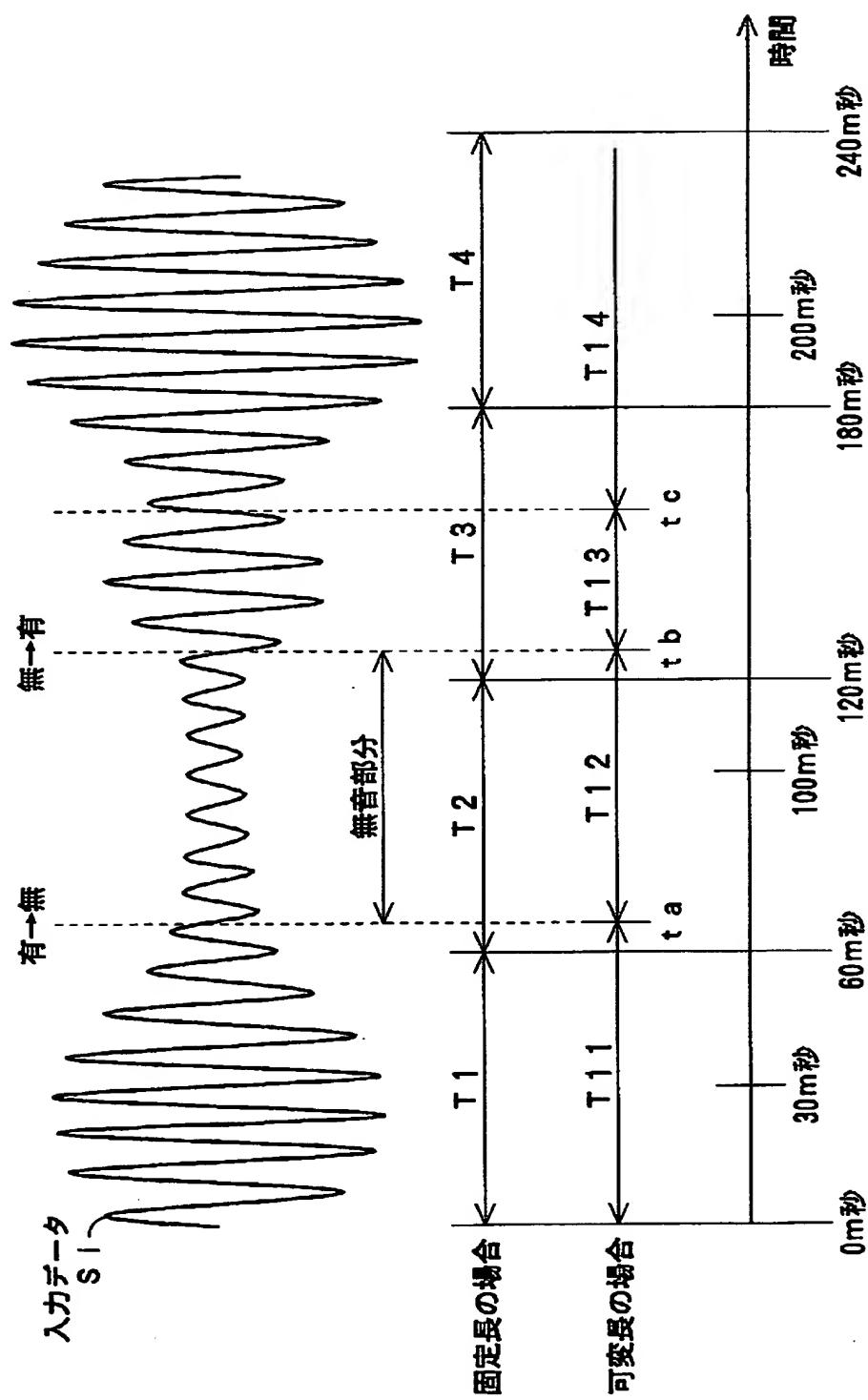


【図 1 5】

5.0 音声レベル対応処理単位期間設定処理ルーチン



【図 1 6】



【書類名】 要約書

【要約】

【課題】 記録媒体から、これに記録された音声信号を定常速度より速い速度で再生し、定常速度に変換して出力する方法で、音声が続切れず、内容を理解できるようにする。

【解決手段】 元の再生音声信号である入力データ  $S_i$  を 60 m 秒 (2880 サンプル) の処理単位期間  $T_1, T_2 \dots$  ごとに区切る。期間  $T_1$  では、有音部分と無音部分が 30 m 秒分ずつ存在し、出力バッファにデータが全く蓄積されていないので、無音部分を全て削除し、前後の有音部分を接合して、有音部分のデータのみを全て出力バッファに書き込む。期間  $T_2$  では、60 m 秒分のデータが全て有音部分であり、出力バッファに 60 m 秒分のデータを全て蓄えるだけの空きがあるので、有音部分のデータをそのまま出力バッファに書き込む。期間  $T_2$  で有音部分が 30 m 秒分未満しか存在しないときには、有音部分の全てに無音部分の一部を付加して、30 m 秒分のデータを出力バッファに書き込む。期間  $T_5$  では、60 m 秒分のデータが全て有音部分であり、出力バッファに 60 m 秒分のデータを全て蓄えるだけの空きがないので、60 m 秒分の有音部分のデータを 30 m 秒分に圧縮して出力バッファに書き込む。

【選択図】 図 3

出 願 人 履 歴 情 報

識別番号 [ 0 0 0 0 0 2 1 8 5 ]

1. 変更年月日	1 9 9 0 年 8 月 3 0 日
[ 変更理由 ]	新規登録
住 所	東京都品川区北品川 6 丁目 7 番 3 5 号
氏 名	ソニー株式会社